Cataract Detection Using Convolutional Neural Network with VGG-19 Model

Md. Sajjad Mahmud Khan
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
sajjad0mahmud@gmail.com

Mahiuddin Ahmed
Department of Computer Science
American International University Bangladesh
Dhaka, Bangladesh
mohiuddin06617@gmail.com

Raseduz Zaman Rasel
Department of Computer Science and Engineering
Dhaka International University
Dhaka, Bangladesh
rasel06@gmail.com

Dr. Mohammad Monirujjaman Khan
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh
monirujjaman.khan@northsouth.edu

***Abstract*—Cataract is one of the prevalent causes of visual impairment and blindness worldwide. There is around 50% of overall blindness. Therefore, an early detection and prevention of cataract may reduce the visual impairment and the blindness. The advancement of Artificial Intelligence (AI) in the field of ophthalmology such as glaucoma, macular degeneration, diabetic retinopathy, corneal conditions, age related eye diseases is quite fruitful unlike cataract. Most of the existing approaches on cataract detection are based on traditional machine learning methods. On the other hand, the manual extraction of retinal features is a time-consuming process and requires an expert ophthalmologist. So, we proposed a model VGG19 which is a convolutional neural network model to detect the cataract by using color fundus images.**

Keywords—Cataract Detection, Deep Learning, VGG19, Convolutional Neural Network

I. Introduction

WHO appraises that about 285 million people are visually impaired globally, with 39 million blinds and 246 million are having moderate to frightful blindness (MSVI) [1].

The current healthcare system is far from satisfactory for the management of common eye diseases due to inadequate medical capabilities especially in middle and low income countries. About 90% of the people who are affected by partial or fully blindness belong to developing countries. However about 75% of vision loss are curable which implies approximately four out of five cases [2]. The most ordinary eye diseases are Glaucoma, Cataract, Diabetic retinopathy (DR), Age-related macular degeneration (AMD), Retinitis pigmentosa, Pterygium and Ocular Surface Neoplasia. Most ocular diseases affect both eyes. However, if treated in early stages 80% of visual impairment are preventable or curable [1]. Late stages of ocular pathologies always lead to severe damage on visual acuity and may be irreversible.

Nevertheless, early screening is not ensured due to the lack of ophthalmologists where important waiting times are registered specially in industrialized countries. Moreover, patient mobility is a limiting factor in particularly aging patients. Thus, an active effort is being made to create and develop methods to automate retinal diseases screening. Many CAD systems have been expanded and are widely used for diagnosing ocular diseases [3].

However, most of the ophthalmologists are usually flooded with general eye check-ups. As a result, ophthalmologists have little time to conduct retinal disease related surgeries. Sufficient number of surgeries are required to prevent blindness due to cataract by experienced ophthalmologists. It's still a challenge to conduct the necessary number of cataract surgeries with the available ophthalmologists. Major challenges for blindness prevention are the public awareness, limited accessibility, high cost of treatment and poor surgical outcomes [2,4].

The main reasons of blindness are Cataract (51%), Glaucoma (8%), Age Related Macular Degeneration (AMD) (5%), Corneal Opacity (4%), Uncorrected Refractive Error (URE) (3%), Trachoma (3%), Diabetic Retinopathy (DR) (1%), and others (25%) [5,6]. The above causes of blindness statistic is represented in Fig. 1 below

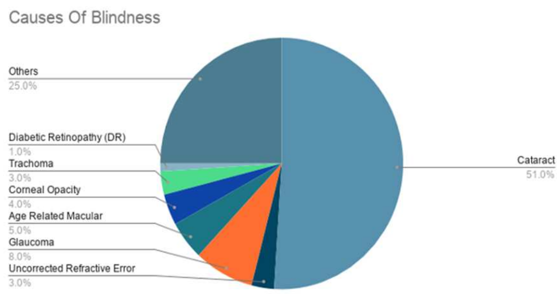


Fig. 1. Causes of Blindness Pie Chart

In ophthalmology, enormous amounts of eye image data and patient related data are available and generated each day. In this field, automated detection of age-related eye conditions such as diabetic retinopathy (DR), age-related macular degeneration and glaucoma had shown impressive outcomes [7,8,9].

Influenced by above mentioned factors we have tried to develop an automated cataract detection system utilizing convolutional neural networks with transfer learning approach. An early detection and proper guided diagnosis could tremendously impact in minimizing the rates of cataracts as automated cataract systems would free up ophthalmologist's time. Though due to lack of proper dataset we were not able to train and develop a state-of-the-art detection system.

The analysis of fundus images for cataract detection have attracted many researchers in the past few years. Over those years multiple automatic cataract detection systems have been developed but these systems could not bring sufficient results due to lack of feature extraction, proper real-world dataset and pre-processing [10,11]. While many of them succeeded in detecting cataracts they either failed in utilizing it on a large scale or had higher requirements to be made. However, the

2021 IEEE World AI IoT Congress (AIIoT) | 978-1-6654-3568-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/AIIOT52608.2021.9454244

978-1-6654-3568-0/21/$31.00 ©2021 IEEE

fundus image analysis for cataract diagnosis is the least focused area of research.

Deep learning (DL) is a subfield of machine learning. DL uses multiple layers to extract higher levels of features from raw input. Because of its feature extraction ability, deep learning is highly used in medical image analysis. To be more specific Convolutional Neural Networks (CNN) are used for medical image analysis. Without the need for human interference, the CNN model extracts local features directly from the fundus images. In [12] a DL based system was implemented on top of Caffe Deep learning framework and cataract was graded with the help of SoftMax function. Zhang et al. [13] implemented a four-level classification network where they utilized deep convolutional neural network and SoftMax function. This DCNN classification network consists of 8 layers. Among them five are convolutional and three are fully connected layers. In [14] a convolutional and deconvolutional network-based cataract classification was performed where CNN was used for fundus images feature extraction and feature learning.

For cataract detection and classification deep learning models have shown optimum performance [15]. However, training of these models need larger retinal datasets to get better performance and the process is very time consuming [15]. Thus, the concepts of transfer learning have been employed in this study to solve such issues.

In [16], based on a decision tree algorithm and 2D Gaussian filter a cataract detection system has been developed by training 1355 retinal fundus images and the accuracy of the developed system was 92.8%. Though decision tree is not capable of dealing with high dimension images.

In this paper an automatic cataract detection system has been proposed based on CNNs which can detect the cataract and non-cataract from retinal fundus images.

II. METHODOLOGY

An automatic cataract detection system is proposed in this study process. The proposed method of cataract detection divided into two different section: pre-processing and DCNN classifier. We will dive deep into above mentioned two sections in this section.

A. Dataset

In this proposed system the used dataset comprises of 800 fundus images of 800 patients. The dataset collected by Shanggong Medical Technology Co., Ltd. from various medical centres in China. The dataset consists of real-life set of patient information. Various cameras were used to capture fundus images in these institutions which result in varied image resolutions. For our purpose we only utilized cataract and normal fundus images from these datasets. The Fig. 2 below illustrates a sample image of cataract vs normal image and Fig. 3 below illustrates the ratio of the normal and cataract image of the dataset.

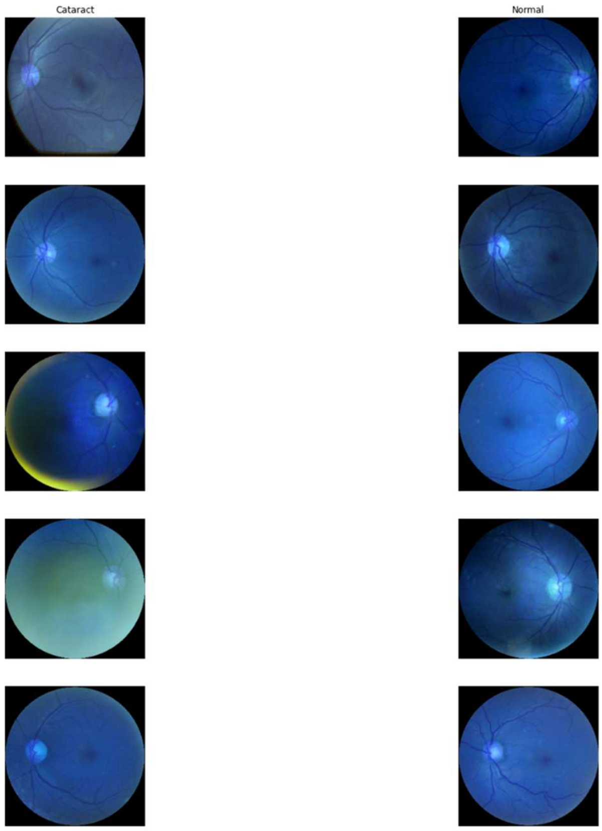


Fig. 2. Cataract vs Normal Image

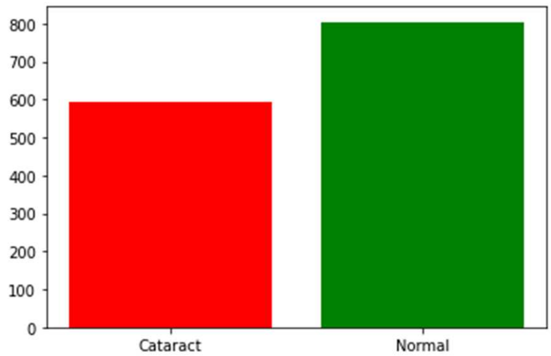


Fig. 3. Cataract and Normal Fundus Image Ratio

The dataset is publicly available at Kaggle [17].

B. Pre-processing

The proposed system dataset consists of images of Normal, Diabetes, Glaucoma, Cataract, Pathological Myopia, Hypertension, Age related Macular Degeneration, other Diseases/Abnormalities. So, in our first step we have filtered out all fundus images other than Cataract and Normal Fundus images. For filtered out data by labels. Experimental fundus images consist of varied image sizes as they were captured using different cameras. So, we have resized image to 224*224 pixels with the help of OpenCV. Then we load and convert the dataset into an array format with the help of NumPy library for training purposes.

C. Convolutional Neural Network

Convolutional Neural Network has become a hot research area in the field of image recognition. CNN has a weight sharing network and this network reduces the complexity of neural networks. This weight sharing network is similar to biological neural networks. However, training a Convolutional Neural Network from scratch is very time consuming.

In computer vision, transfer learning is a hot topic as it allows to build model with high accuracy in a convenient way. Expeditious model building is possible because of pre-trained models. Pre-trained models are model which is trained with millions of large benchmark dataset. Thus, pre-trained are able to deliver faster training time. Additionally, Fine-tuning CNN networks are much simpler and faster than training a network from scratch [18,19].

VGGNet abbreviated as visual geometry group network is a multi-layer operated convolutional neural network. VGGNet is a CNN model and is trained on the ImageNet dataset. ImageNet, is a dataset of 15 million labelled high-resolution images. Instead of larger filters, VGGNet uses 3x3 filter to increase model depth level. A variation of VGGNet named VGG-19 has 19 layers including convolutional, fully connected, max pooling and dropout layers. Max pooling layers is used to reduce the volume size in VGG-19. Each of the convolutional layer is stacked with an activation function and max-pooling operation. VGGNet consists of two fully connected layers with 4096 channels. Above mentioned each layer is followed by another fully connected layer with 1000 channels to predict 1000 labels. To reduce feature dimensionality some of convolutional layers have been associated with max pooling layers. First two layers are convolutional layers which use 64 filters where filter size is 3x3. For feature extraction purpose convolutional layers filters were applied on the input images. To prepare feature vector fully connected layers were used.

As we were not able to utilize large dataset, we have chosen to use VGG-19 pre-trained model. As VGG-19 model is availed as feature extractor for this study, only full connected layers will be trained. Fig. 4 below illustrates our proposed systems VGG19 model.

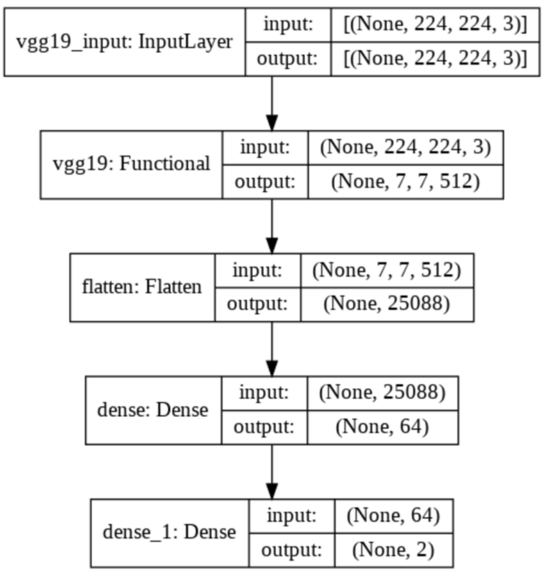


Fig. 4. Proposed Systems VGG-19 Model

III. Training

The proposed convolutional neural network takes 224 × 224px processed RGB image as input and generates the output which is classification of cataract. For pre-trained VGG-19 model only classification part needs to be trained which consists of dense and dropout layers. We applied VGG-19 model on our training dataset and output vectors last layer of this employment was used as input in detection part. As we are using the pre-trained VGG-19 model we only need to train the classification part. Classification contains only dense and dropout layers. We applied VGG-19 model on our training dataset. We took output vectors from the last layer as the training input for the detection part.

Additionally, SoftMax and ReLU activation functions have been applied to learn complex patterns from the fundus images.

We have chosen to use the Adam optimization algorithm as optimization algorithm which is a replacement of the stochastic gradient descent (SGD). In computer vision and natural language processing, ADAM has recently seen broader adoption as it helps to reduce cost function thus producing more useful models. Optimization algorithms attempt to obtain optimal weights, mitigate errors and enhance accuracy in a memory efficient way. Using Backpropagation, the optimizer is responsible for updating the neuron weights. It calculates the partial derivative of the loss function with respect to weights and the weights are subtracted from it. Adam optimization algorithm calculates an exponential weighted moving average of the gradient and then squares the calculated gradient. In each iteration step size of Adam is more or less bounded by the step size hyperparameter. Furthermore, Adam optimization algorithm is invariant to the diagonal rescale of the gradient.

IV. Results

In this section, this study's evaluation criteria and experimental results are represented. Additionally, some details of the dataset are explained.

A. *Evaluation Criteria*

In this study, to measure the performance of cataract detection model Accuracy, Precision and Recall were employed. The formulas of them are

$$\mathrm{X_a} = \frac{\mathrm{T_p} + \mathrm{T_n}}{\mathrm{T_p} + \mathrm{F_p} + \mathrm{T_n} + \mathrm{F_n}}$$

$$X_p = \frac{\mathrm{T}_p}{\mathrm{T}_p + \mathrm{F}_p}$$

$$X_r = \frac{\mathrm{T}_p}{\mathrm{T}_p + \mathrm{F}_p}$$

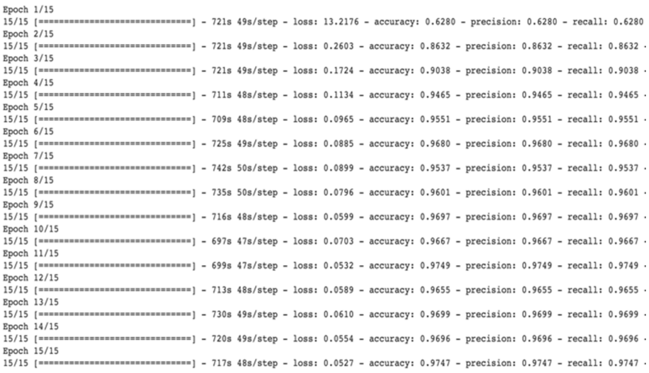
Where X_a, X_p, X_r represents accuracy, precision and recall respectively and T_p, T_n, F_p, and F_n are the number of true positive, true negative, false positive, and false negative samples.

B. *Experimental results*

In this section, proposed systems experimental results are described. Proposed system is implemented on top of pre-trained model VGGNet. VGGNet's variation VGG-19 is utilized for this system. KAGGLE datasets were used for the experiments training and testing. To achieve significant results with a flexible implementation of VGGNet, general purpose units (GPUs) were used. For an unbiased evaluation, we have split the dataset into training and testing randomly. For splitting we have applied cross validation following 80/20 rule. On these cross-validated data 80% were put into training and 20% were put into testing. By testing on this dataset our model achieved below results in Table 1 and Fig. 5 illustrates the image of the results of the training.

TABLE 1 Experimental Results

Accuracy(%)	*Precision(%)*	*Loss(%)*
97.47	97.47	5.27



```
Epoch 1/15
15/15 [==============================] - 721s 49s/step - loss: 13.2176 - accuracy: 0.6280 - precision: 0.6280 - recall: 0.6280
Epoch 2/15
15/15 [==============================] - 721s 49s/step - loss: 0.2603 - accuracy: 0.8632 - precision: 0.8632 - recall: 0.8632 -
Epoch 3/15
15/15 [==============================] - 721s 49s/step - loss: 0.1724 - accuracy: 0.9038 - precision: 0.9038 - recall: 0.9038 -
Epoch 4/15
15/15 [==============================] - 711s 48s/step - loss: 0.1134 - accuracy: 0.9465 - precision: 0.9465 - recall: 0.9465 -
Epoch 5/15
15/15 [==============================] - 709s 48s/step - loss: 0.0965 - accuracy: 0.9551 - precision: 0.9551 - recall: 0.9551 -
Epoch 6/15
15/15 [==============================] - 725s 49s/step - loss: 0.0885 - accuracy: 0.9680 - precision: 0.9680 - recall: 0.9680 -
Epoch 7/15
15/15 [==============================] - 742s 50s/step - loss: 0.0899 - accuracy: 0.9537 - precision: 0.9537 - recall: 0.9537 -
Epoch 8/15
15/15 [==============================] - 735s 50s/step - loss: 0.0796 - accuracy: 0.9601 - precision: 0.9601 - recall: 0.9601 -
Epoch 9/15
15/15 [==============================] - 716s 48s/step - loss: 0.0599 - accuracy: 0.9697 - precision: 0.9697 - recall: 0.9697 -
Epoch 10/15
15/15 [==============================] - 697s 47s/step - loss: 0.0703 - accuracy: 0.9667 - precision: 0.9667 - recall: 0.9667 -
Epoch 11/15
15/15 [==============================] - 699s 47s/step - loss: 0.0532 - accuracy: 0.9749 - precision: 0.9749 - recall: 0.9749 -
Epoch 12/15
15/15 [==============================] - 713s 48s/step - loss: 0.0589 - accuracy: 0.9655 - precision: 0.9655 - recall: 0.9655 -
Epoch 13/15
15/15 [==============================] - 730s 49s/step - loss: 0.0610 - accuracy: 0.9699 - precision: 0.9699 - recall: 0.9699 -
Epoch 14/15
15/15 [==============================] - 720s 49s/step - loss: 0.0554 - accuracy: 0.9696 - precision: 0.9696 - recall: 0.9696 -
Epoch 15/15
15/15 [==============================] - 717s 48s/step - loss: 0.0527 - accuracy: 0.9747 - precision: 0.9747 - recall: 0.9747 -
```

Fig. 5. Training Image

V. Conclusion

In this paper, an automatic cataract detection system is proposed where pre-trained model VGGNet was employed. To be more specific VGG-19 were used for this study which achieved an accuracy of 97.47% on the test dataset. Dataset as mentioned above are collected from KAGGLE. This result shows that presented method achieve high accuracy even on unfiltered and image quality unassessed fundus photographs without ophthalmologist's intervention. However, better and more in-depth researched feature extraction techniques along with classification algorithms are essential. Such solutions would be able to increase accessibility to healthcare, reduce screening cost and time for both ophthalmologist and patient and early diagnosis. Considering above reasoning, automated cataract detection system would be really helpful for developing countries where trained ophthalmologist against patient ratio is inadequate. In future, we would like to study this presented system upon large dataset, evaluate this proposed system with other deep learning architectures, scale up this system for diabetic retinopathy and glaucoma detection. Additionally, due to shortage of proper resources and a dataset we were not able to develop cataract severity grading and exact subtle lesion detection system. Cataract severity and subtle lesion detection is also left for future development.

References

[1] S. Flaxman et al., "Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis", The Lancet Global Health, vol. 5, no. 12, pp. e1221-e1234, 2017. Available: 10.1016/s2214-109x(17)30393-5.

[2] World Health Organization. 2013. Universal eye health: a global action plan 2014-2019. [online] Available at: <https://www.who.int/blindness/actionplan/en/> [Accessed 1 February 2021].

[3] Z. Zhang et al., "A survey on computer aided diagnosis for ocular diseases", BMC Medical Informatics and Decision Making, vol. 14, no. 1, 2014. Available: 10.1186/1472-6947-14-80.

[4] M. Deshpande, "Vision 2020: Right to Sight - India", Medical Journal Armed Forces India, vol. 64, no. 4, pp. 302-303, 2008. Available: 10.1016/s0377-1237(08)80002-0.

[5] D. Pascolini and S. Mariotti, "Global estimates of visual impairment: 2010", 2021.

[6] S. Flaxman et al., "Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis", The Lancet Global Health, vol. 5, no. 12, pp. e1221-e1234, 2017.

[7] D. Ting et al., "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multi-ethnic Populations With Diabetes", JAMA, vol. 318, no. 22, p. 2211, 2017. Available: 10.1001/jama.2017.18152.

[8] V. Gulshan et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs", JAMA, vol. 316, no. 22, p. 2402, 2016. Available: 10.1001/jama.2016.17216.

[9] E. Pead et al., "Automated detection of age-related macular degeneration in colour fundus photography: a systematic review", Survey of Ophthalmology, vol. 64, no. 4, pp. 498-511, 2019. Available: 10.1016/j.survophthal.2019.02.003.

[10] L. Guo, J. Yang, L. Peng, J. Li and Q. Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis", Computers in Industry, vol. 69, pp. 72-80, 2015. Available: 10.1016/j.compind.2014.09.005.

[11] J. Yang et al., "Exploiting ensemble learning for automatic cataract detection and grading", Computer Methods and Programs in Biomedicine, vol. 124, pp. 45-57, 2016. Available: 10.1016/j.cmpb.2015.10.007.

[12] Y. Dong, Q. Zhang, Z. Qiao and J. Yang, "Classification of cataract fundus image based on deep learning", 2017 IEEE International Conference on Imaging Systems and Techniques (IST), 2017. Available: 10.1109/ist.2017.8261463.

[13] Linglin Zhang et al., "Automatic cataract detection and grading using Deep Convolutional Neural Network", 2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC), 2017. Available: 10.1109/icnsc.2017.8000068.

[14] X. Xu, L. Zhang, J. Li, Y. Guan and L. Zhang, "A Hybrid Global-Local Representation CNN Model for Automatic Cataract Grading", IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 2, pp. 556-567, 2020. Available: 10.1109/jbhi.2019.2914690.

[15] A. Imran, J. Li, Y. Pei, J. Yang and Q. Wang, "Comparative Analysis of Vessel Segmentation Techniques in Retinal Images", IEEE Access, vol. 7, pp. 114862-114887, 2019. Available: 10.1109/access.2019.2935912

[16] L. Xiong, H. Li and L. Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis", Journal of Healthcare Engineering, vol. 2017, pp. 1-16, 2017. Available: 10.1155/2017/5645498.

[17] "Ocular Disease Recognition", Kaggle.com, 2021. [Online]. Available: https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k. [Accessed 11 February 2021].

[18] B. Huynh, H. Li and M. Giger, "Digital mammographic tumor classification using transfer learning from deep convolutional neural networks", Journal of Medical Imaging, vol. 3, no. 3, p. 034501, 2016. Available: 10.1117/1.jmi.3.3.034501

[19] H. Chen et al., "Standard Plane Localization in Fetal Ultrasound via Domain Transferred Deep Neural Networks", IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 5, pp. 1627-1636, 2015. Available: 10.1109/jbhi.2015.2425041